Super Resolution results in PANOPTES, an adaptive multi-aperture folded architecture
— Source link [↗]

Esmaeil Faramarzi, Vikrant R. Bhakta, Dinesh Rajan, Marc P. Christensen

Institutions: Southern Methodist University

Related papers:

- Pixel super resolution imaging method based on coded aperture modulation

- Spaceborne multispectral imaging system design method based on super-resolution reconstruction

- Spaceborne super-resolution imaging system design method

- An efficient iterative super-resolution technology for coded aperture imaging

- Single-molecule positioning based fast super-resolution imaging method and system

SUPER RESOLUTION RESULTS IN PANOPTES, AN ADAPTIVE MULTI-APERTURE FOLDED ARCHITECTURE

Esmaeil Faramarzi, Vikrant R. Bhakta, Dinesh Rajan, and Marc P. Christensen

Department of Electrical Engineering, Southern Methodist University, 6251 Airline Road, Dallas, Texas 75275-0338
efaramarzi, vrbhakta, rajand, mpc @lyle.smu.edu

ABSTRACT

We present experimental results of digital super resolution (DSR) techniques on low resolution data collected using PANOPTES, a multi-aperture miniature folded imaging architecture. The flat form factor of **PANOPTES** architecture results in an optical system that is heavily blurred with space variant PSF which makes super resolution challenging. We also introduce a new DSR method called SRUM (Super-Resolution with Unsharpenning Mask) which can efficiently highlight edges by embedding an unsharpenning mask to the cost function. This has much better effect than just applying the mask after all iterations as a post-processing step.

1. INTRODUCTION

Capturing high quality images and videos requires high-cost and bulky optical elements whose physical sizes dictate the light-gathering capability and the resolving power of the imaging system, a constraint that has persisted since their invention [1]. In recent years there has been a big trend toward generating high resolution (HR) images/videos by fusing information from a series of low resolution (LR) image/video frames. This trend has activities in two main research areas, multiplexed computational imaging (MCI) and digital super resolution (DSR) [2]. MCI primarily focuses on joint design of optics, detector and reconstruction algorithm. On the other hand, DSR deals with software techniques for reconstructing a HR image from a set of blurry, undersampled and noisy LR images.

Most DSR techniques in the literature are designed to operate on LR images having motion in their contents due to the camera or scene movement. These techniques employ the phase variation between the frequency spectrums of captured images to remove aliasing. By contrast, motion-free DSR techniques use variations in the amplitude of the transfer function to make resolution enhancement. In this paper, we focus on the methods that rely on variations in the phase of DSR system transfer function resulting from inter-frame subpixel shifts.

A multi-aperture imaging system architecture called **PANOPTES** (Processing Arrays of Nyquist-limited Observations to Produce a Thin Electro-optic Sensor) was proposed [1] and consists of several tiled sub-imagers, each with a built-in steering mechanism. **PANOPTES** incorporates intelligent image acquisition adaptability and real-time resource allocation capability based upon the information content of the scene. In [3] an unobscured, folded architecture that utilizes off-axis tilted paraboloid as a sole power surface was demonstrated for the first time ever at a small scale ($8mm \times 6mm \times 5mm$), as shown in Fig. 1, but due to its unique optics the system is heavily blurred and has space variant PSF that makes it challenging to use existing DSR algorithms. Experimental validation integration capabilities of **PANOPTES** architecture were demonstrated in [3].
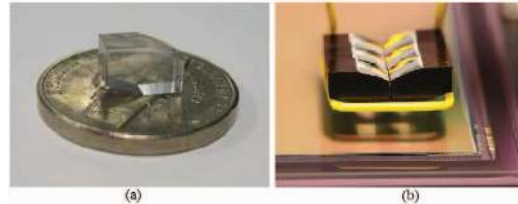

Fig. 1. Photographs of (a) a single uncoated **PANOPTES** sub-imager and (b) a tiled array of **PANOPTES** sub-imagers arranged on a large CCD detector.

In an imaging system when the detector Nyquist cutoff frequency is appreciably lower than the optical cutoff frequency, there exists observable amount of aliasing in the captured images. Aliased components contain valuable high frequency information and their existence guarantees the possibility of getting enhancement by super resolution. Fig. 2 demonstrates the typical DSR regime from the perspective of the optical and detector modulation transfer function (MTF) plots. It is clear from this figure that the detector-limited nature of imaging system results in aliasing. DSR allows recovery of this aliased information and in principle it is possible to capture image data up to the optical cutoff from multiple sub-pixel shifted low-resolution images. Hence it becomes important to identify the region in the image plane where the retrievable optical spatial frequencies exceed the detector Nyquist cutoff. This region is termed the sweet spot of the **PANOPTES** sub-imager which is a patch in the center of the captured images.

While there are some papers on restoring a single image degraded by spatially variant blur (e.g. [4]) and also a few works on DSR for space variant images (e.g.[5]), to the best of our knowledge no paper has shown its results on real space variant images captured by a flat camera. In this paper we illustrate the reconstruction results of several well-known DSR techniques in order to validate the proposed mechanism in **PANOPTES** architecture. We first show the results of these techniques on a synthetic data and then on a central image region surrounding the sweet spot in an image set containing 121 shifted images. These images are captured by a QI-Cam 4000R charge-coupled device (CCD) scientific camera detector with F/# of 2 and detector pixel pitch of 7.4 μm [3]. Also we illustrate a new method called SRUM (Super-Resolution with Unsharpenning Mask) which results HR images with higher contrast than some other well-known DSR methods through embedding an unsharpenning mask into the iteration process.

This paper is organized as follows: Section 2 explains the most common forward model used to formulate DSR methods. Section 3 describes a few DSR techniques. In section 4 we introduce the SRUM method. Finally section 5 depicts the experimental results of the SRUM and other methods on both synthetic and real image sets.
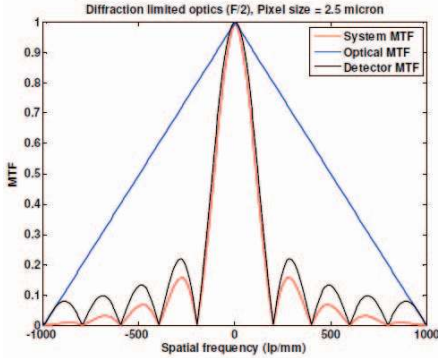
Fig. 2. An example of digital super-resolution (DSR) regime. DSR allows recovery of aliased information, caused by detector limited performance of an imaging system, up to the optical cutoff frequency using sub-pixel shifted observations.

2. DSR FORWARD MODEL

The linear forward imaging model in the spatial domain which illustrates the process of generating LR images from a HR image can be defined as:

$$g_k(x,y) = \left(f(\acute{x}, \acute{y}) * h_k(x,y)\right) \downarrow_L + n_k(x,y), \qquad k = 1, \dots, P$$

$$\begin{bmatrix} \acute{x} \\ \acute{y} \end{bmatrix} = z_k \begin{bmatrix} cos(\theta_k) & -sin(\theta_k) \\ sin(\theta_k) & cos(\theta_k) \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} s_{xk} \\ s_{yk} \end{bmatrix} \qquad (1)$$

where f is the original HR image of size $N_1 \times N_2$ and g_k is the kth of total P observed LR images of size $M_1 \times M_2$. According to this model, the HR image is first zoomed, rotated and shifted horizontally and vertically by the amounts of z_k, θ_k, s_{xk} and s_{yk} respectively, then convolved with the point spread function (PSF) of the lens and other blurring effects such as motion, then downsampled by a factor of L, and finally corrupted by noise to get the LR image. We have: $N_1 = LM_1$ and $N_2 = LM_2$. In matrix notation, (1) can be written as:

$$\mathbf{g}_k = \mathbf{W}_k \mathbf{f} + \mathbf{n}_k, \qquad k = 1, \dots, P$$
$$\mathbf{W}_k = \mathbf{D}_k \mathbf{H}_k \mathbf{S}_k \qquad (2)$$

where $\mathbf{f}$ is the input image in lexicographical notation resulting in a vector of size $N_1 N_2 \times 1$, $\mathbf{S}_k$ and $\mathbf{H}_k$ are kth warping and blurring operators of size $N_1 N_2 \times N_1 N_2$, $\mathbf{D}_k$ is the downsampling matrix of size $M_1 M_2 \times N_1 N_2$, and $\mathbf{g}_k$ and $\mathbf{n}_k$ are the vectors of kth LR image and noise respectively, both of size $M_1 M_2 \times 1$. Typically $\mathbf{D}_k$ and $\mathbf{H}_k$ are the same for all LR images. Alternatively, the model in (2) can be expressed in terms of the entire set of LR images as:

$$\mathbf{f} = \mathbf{Wf} + \mathbf{n} \qquad (3)$$

where $\mathbf{g} = \left[\mathbf{g}_1^T, \mathbf{g}_2^T, \dots, \mathbf{g}_P^T\right]^T$ and $\mathbf{n} = \left[\mathbf{n}_1^T, \mathbf{n}_2^T, \dots, \mathbf{n}_P^T\right]^T$ are vectors of size $PM_1 M_2 \times 1$ and $\mathbf{W} = \left[\mathbf{W}_1^T, \mathbf{W}_2^T, \dots \mathbf{W}_P^T\right]^T$ is a matrix of size $PM_1 M_2 \times N_1 N_2$.

The estimation of motion information, which is referred to as image registration, is of great importance in super resolution and its precision has a direct influence on the reconstruction quality.

3. OVERVIEW OF DSR TECHNIQUES

Irani and Peleg [6] proposed an iterative back-projection (IBP) approach by the use of steepest decent optimization technique. First an initial estimate of the HR image is calculated using a simple method like interpolation. Then by simulating the imaging process, the estimated HR image is projected onto the LR image plane to obtain a set of LR images corresponding to the observed

LR input images. Finally the differences between the observed and estimated LR images are back-projected onto the HR image plane to update the HR image. This process is repeated iteratively. So we have:

$$\hat{\mathbf{f}}^{n+1} = \hat{\mathbf{f}}^n - \lambda/2 \; \mathbf{\nabla}^n$$
$$\mathbf{\nabla}^n = \nabla J^n = 2\mathbf{W}^T\left(\mathbf{W}\hat{\mathbf{f}}^n - \mathbf{g}\right) \qquad (4)$$

Here $\hat{\mathbf{f}}^n$ is the estimated HR image at the iteration n, $\mathbf{\nabla}^n$ is the gradient of the cost function J (refer to (6)) at the iteration n, and λ called step size is a constant that dictates the strength of the updating. This method is basically derived based upon minimizing a least-squares (LS) cost function which minimizes the L2 norm of the residual vector:

$$J(\mathbf{f}) = \|\mathbf{Wf} - \mathbf{g}\|_2^2 = (\mathbf{Wf} - \mathbf{g})^T (\mathbf{Wf} - \mathbf{g}) = \mathbf{f}^T \mathbf{W}^T \mathbf{Wf} - \mathbf{f}^T \mathbf{W}^T \mathbf{g} - \mathbf{g}^T \mathbf{Wf} + \mathbf{g}^T \mathbf{g} = \mathbf{f}^T \mathbf{W}^T \mathbf{Wf} - 2\mathbf{f}^T \mathbf{W}^T \mathbf{g} + \mathbf{g}^T \mathbf{g}$$

$$\hat{\mathbf{f}} = \frac{\text{argmin}}{\mathbf{f}} J(\mathbf{f}) = \nabla J(\mathbf{f}) = \partial J(\mathbf{f})/\partial \mathbf{f} = 2\mathbf{W}^T(\mathbf{Wf} - \mathbf{g}) \qquad (5)$$

Various modifications of this method have appeared in the literature, depending on the type of norm function, existence or nonexistence of a regularization term, and steadiness or variability of the step size parameter λ. Most often a regularization term is added to the cost function in order to more stabilize the inverse problem of estimating F from G:

$$\hat{\mathbf{f}} = \text{argmin } J(\mathbf{f}) = \text{argmin } \|\mathbf{Wf} - \mathbf{g}\|_2^2 + \gamma\rho(\mathbf{f}) \qquad (6)$$

Hardie et al. [7] proposed a maximum a posteriori (MAP) framework for jointly estimating the high-resolution image and the registration parameters by considering noise and priors as Gaussian random variables. The cost function used is again the minimum of least-square error and the iteration algorithm is steepest descent. At each iteration the best updating parameter is found through minimizing the cost function. Then in [8] they updated their work by using the Keren registration method [9] instead of estimating and updating the registration parameters at the iterations.

Elad and Feuer [10] introduced using LS cost function with L2-norm for reconstructing of continuous image sequences from the view point of adaptive filtering theory. The cost function here is a function of time and the assumed correlation between the continuous movie frames is applied to simplify the equations.

Farsiu et al. [11] proposed a robust method by using L1-norm instead of L-2 norm in both the measurement and the regularization terms. They showed that the L1-norm converges to median estimation which has the highest breakpoint value. In this work, an edge preserving regularization term called Bilateral-TV is used.

In [2] El-Yamani and Papamichalis suggested applying the Lorentzian error norm to better suppress the effect of large errors in the cost function.

We provide a comparison of reconstruction from selected methods in Section 5.

4. SRUM METHOD

In this section, we introduce a new super-resolution method called SRUM. By adding an unsharpenning mask to the IBP method explained in the previous section, we can achieve a reconstruction result that has a better visual perception for many sorts of images through highlighting edges and details. If $f(x,y)$ is an image, then:

$$g(x,y) = f(x,y) + \beta f(x,y) * h_{hp}(x,y) \qquad (7)$$

will be an image with more emphasis on the edges [12]. Herein β is a constant and h_{hp} is a high-pass filter. The first term is added to prevent loss of the background tonality caused by the second term because the highpass filter zeros out the dc term. To use this technique with the IBP method, we embed a matrix $\mathbf{H}$ into the cost

function in (6). We also add a regularization term to control the smoothness of the reconstructed image:

$$J(\mathbf{f}) = \mathbf{f}^T\mathbf{W}^T\mathbf{H}\mathbf{W}\mathbf{f} - 2\mathbf{f}^T\mathbf{W}^T\mathbf{g} + \mathbf{g}^T\mathbf{g} + \gamma\mathbf{f}^T\mathbf{L}^T\mathbf{L}\mathbf{f}$$
$$\mathbf{H} = \mathbf{I} + \beta\mathbf{H}_{hp} \tag{8}$$

where $\mathbf{I}$ is the identity matrix, and $\mathbf{H}_{hp}$ and $\mathbf{L}$ are the convolution kernels related to the filter $h_{hp}(x, y)$ in (7) and another highpass filter $l(x, y)$, respectively. One choice for these filters is the Laplacian operator which is a second derivative mask with some choices as below:

$$\begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix}, \begin{bmatrix} 1/2 & 0 & 1/2 \\ 0 & -2 & 0 \\ 1/2 & 0 & 1/2 \end{bmatrix}, \begin{bmatrix} 1/6 & 2/3 & 1/6 \\ 2/3 & -10/3 & 2/3 \\ 1/6 & 2/3 & 1/6 \end{bmatrix}$$

We used the third matrix in our implementation. Now the gradient is changed to:

$$\nabla = \nabla J = 2\mathbf{W}^T(\mathbf{H}\mathbf{W}\mathbf{f} - \mathbf{g}) + 2\gamma\mathbf{L}^T\mathbf{L}\mathbf{f} \tag{9}$$

and the iteration equation is replaced by:

$$\hat{\mathbf{f}}^{n+1} = \hat{\mathbf{f}}^n - \lambda^n\left[\mathbf{W}^T(\mathbf{H}\mathbf{W}\hat{\mathbf{f}}^n - \mathbf{g}) + \gamma\mathbf{L}^T\mathbf{L}\hat{\mathbf{f}}^n\right] \tag{10}$$

Hence at each iteration instead of the HR image itself, a more contrasted version of the HR image is projected onto the LR image plane. As we shown in section (5), this has absolutely different effect than just applying the mask after all iterations as a post-processing step.

The optimal value for the step size parameter λ^n at the nth iteration can be found by the following approach:

$$\lambda^n = \underset{\lambda^n}{\arg\min} J(\mathbf{f}^{n+1}) = \underset{\lambda^n}{\arg\min} J\left(\hat{\mathbf{f}}^n - \lambda^n/2\ \nabla^n\right) \tag{11}$$

and after some manipulation, we can evaluate λ^n as:

$$\lambda^n = \frac{\bar{\nabla}^{nT}(\mathbf{H}\mathbf{W}\hat{\mathbf{f}}^n - \mathbf{g}) + \gamma\bar{\bar{\nabla}}^{nT}\mathbf{L}\hat{\mathbf{f}}^n}{\bar{\nabla}^{nT}\bar{\nabla}^n + \gamma\bar{\bar{\nabla}}^{nT}\bar{\bar{\nabla}}^n} \tag{12}$$

where:

$$\bar{\nabla}^n = \mathbf{W}\nabla^n, \quad \bar{\bar{\nabla}}^n = \mathbf{L}\nabla^n \tag{13}$$

Results on the application of synthetic and real data are shown in the next section.

5. RECONSTRUCTION RESULTS OF DIFFERENT ALGORITHMS

In this section we show the results of selected and modified DSR techniques mentioned in the two previous sections on synthetic LR images generated from a HR image through the assumed observation model and also on the real images taken by a flat low quality camera [1][3] in order to validate the PANOPTES system architecture. The methods compared here are MAP [7][8], IBP [6], Robust [11], nonlinear interpolation [13] followed by Wiener filtering and SRUM (section 4).

Fig. 3 depicts DSR reconstruction for a synthetic image set. The ground truth HR image is presented in Fig. 3(a). By means of translational warping, blurring by a Gaussian of variance 1, downsampling by a factor of 4 and adding Gaussian noise of variance 25, 16 LR images like Fig. 3(b) are generated which is zoomed in to the size of HR image through zero order interpolation for better comparison. Fig. 3(c) illustrates the MAP reconstruction result. IBP and Robust methods also produce very similar results to MAP and are slightly better than of nonlinear interpolation method. So by using the exact prior information about warping, blurring and noise, MAP, IBP and Robust methods make very similar results with comparable signal to noise ratios. Also if there is ambiguity in the forward model (e.g. lens PSF, registration parameters, noise information, etc.) that is almost uniform for the entire LR images, all these three DSR methods have similar

behavior. On the other hand, if ambiguity in the forward model exists only in a few LR images, Robust method has a better tolerance because of using L1 norm in its cost function. To show this, we repeat the previous experiment but this time one of LR images has been rotated, zoomed in, blurred with a different PSF and considered with absolutely wrong shift parameters, as shown in Fig. 4(a). The reconstruction results for MAP and Robust methods are presented in Fig. 4(b) and Fig. 4(c), respectively. As seen, the errors in the MAP image are evident while the Robust image is still clean.

Fig. 5 and Fig. 6 show the results of SRUM method with different values of β for the example in Fig. 3 and also for a real image set [14] having 26 LR images. Keren registration method [9] implemented in EPFL software package [15] was employed for registering the images. As seen, the effects of adding the unsharpenning mask are increasing the image contrast and producing sharper edges which give a better visual perception for many kinds of images. Fig. 7 shows a comparison between the effects of embedding this mask into each iteration and adding it as a postprocessing step.

One of the most challenging issues in the *PANOPTES* architecture arises from the need to maintain a flat form factor while simultaneously being able to produce HR imagery. An expected outcome of the fast F/# and the tilted arrangement of the optical elements is that the sub-imager's PSF is large and varies noticeably as a function of field location. Fig. 8(a) represents one of 121 resulted LR images taken for evaluating the *PANOPTES* system. It is obvious from this image that the left side is much blurred than the right side and the farther away along the transverse plane an object point is from the on-axis location, the larger is its PSF. As noted before, it becomes important to identify the sweet spot which is the region in the image plane where the retrievable optical spatial frequencies exceed the detector Nyquist cutoff. Fig. 8(b) depicts the central patch of size 40×40 pixels that includes the sweet spot. Fig. 8(c) shows the reconstruction results of the MAP method for scale factor of 3 (L in equation (1)). No prior information about the PSF and registration parameters is used for reconstruction. Again, the shift values are estimated by Keren method [9]. Fig. Fig. 8(d) illustrates the reconstruction results for the SRUM method which clearly has more contrast than the MAP method.
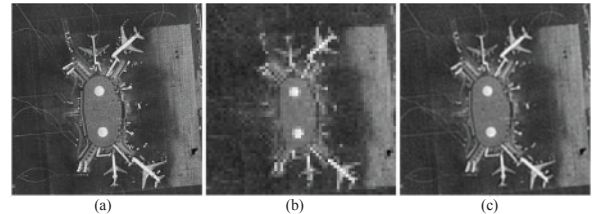
6. ACKNOWLEDGEMENT

Fig. 3. Simulation results of different SR methods. (a) The ground truth HR image. (b) One of 16 LR images. (c) MAP reconstruction result.
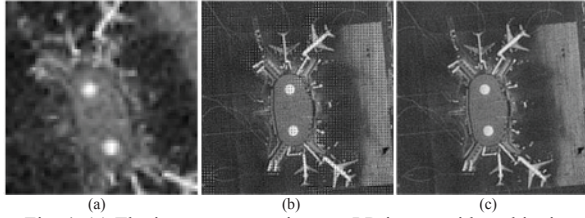
Fig. 4. (a) The image set contains one LR image with ambiguity in its forward model. (b) IBP reconstruction result. (c) Robust reconstruction result.
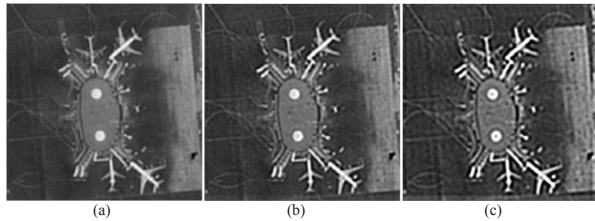

Fig. 5. SRUM reconstruction results with (a) $\beta = \gamma = 0$ (the same as IBP), (b) $\beta = \gamma = 0.15$, and (c) $\beta = \gamma = 0.2$.
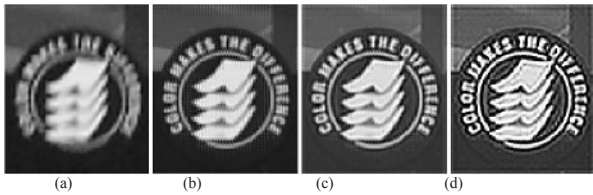

Fig. 6. (a) One LR image in a real image set. IBP results with (b) $\beta = \gamma = 0$, (c) $\beta = \gamma = 0.15$, and (d) $\beta = \gamma = 0.25$.
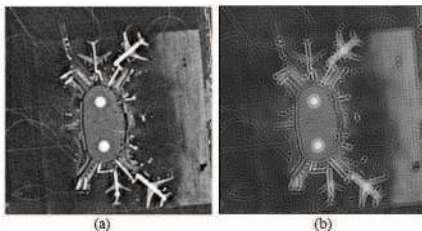

Fig. 7. A comparison between the effects of (a) embedding the unsharpenning mask into each iteration of IBP method and (b) Adding the mask as a postprocessing step.
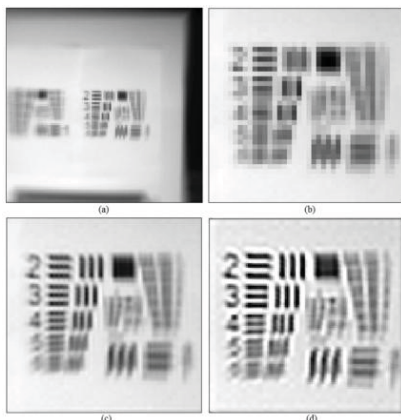

Fig. 8. (a) One of 121 images taken by a flat camera. (b) The region which contains the sweet spot. (c) MAP and (d) SRUM results with DSR scale factor of 3.

7. REFERENCES

[1] M. P. Christensen et al., "Adaptive Flat Multi-resolution Multiplexed Computational Imaging Architecture utilizing Micro-mirror arrays to Steer Sub-imager Fields-of-view," *in Journal of Applied Optics*, Vol. 45, Issue 13, pp. 2883-2892, May 2006.

[2] N. El-Yamany et al., "Adaptive framework for robust high-resolution image reconstruction in multiplexed computational imaging architectures," *Applied Optics*, Vol. 47, No 10, pp. B117-B127, April 2008.

[3] Vikrant R. Bhakta et al., "Experimentally Validated Computational Imaging with Adaptive Multi-aperture Folded Architecture," *Accepted in Feature Issue of Applied Optics*, October 2009.

[4] J. Nagy and D. P. O'Leary, "Restoring images with spatially varying blur", *SIAM J. Scientific Computing*, Vol. 19, pp. 1062–1083, 1998.

[5] D. Rajan, S. Chaudhuri, "Simultaneous Estimation of Super-Resolved Scene and Depth Map from Low Resolution Defocused Observations," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, no. 9, pp. 1102-1117, Sept. 2003

[6] M. Irani and S. Peleg, "Improving resolution by image registration," *CVGIP: Graph. Models Image Processing*, vol. 53, pp. 231–239, 1991.

[7]   R. C. Hardie et al., "Joint MAP Registration and High Resolution Image Estimation Using a Sequence of Undersampled Images," *IEEE Transactions on Image Processing*, Vol. 6, No. 12, Dec. 1997, pp. 1621-1633.

[8] R. C. Hardie et al., "High Resolution Image Reconstruction from a Sequence of Rotated and Translated Frames and its Application to an Infrared Imaging System," *Optical Engineering*, Vol. 37, No. 1, pp. 247-260, Jan. 1998.

[9] D. Keren et al., "Image sequence enhancement using sub-pixel displacement," *in Proceedings IEEE Conference on Computer Vision and Pattern Recognition*, pp. 742-746, June 1988.

[10] M. Elad and A. Feuer, "Super-Resolution Restoration of Continuous Image Sequence - Adaptive Filtering Approach," *IEEE Trans. on Image Processing*, Vol. 8. No. 3, pp. 387-395, March 1999.

[11] S. Farsiu et al., "Fast and Robust Multi-Frame Super-resolution", *IEEE Trans. On Image Processing*, Vol. 13, No. 10, pp. 1327-1344, October 2004.

[12] R. Gonzalez and R. Woods, Digital Image Processing, 3rd ed. (Prentice Hall, 2008).

[13] S. C. Park et al., "Super-resolution image reconstruction: a technical overview," IEEE Signal Processing Magazine, 20(3):21–36, May 2003.

[14] P. Milanfar, MSDP super-resolution and demosaicing datasets, http://users.soe.ucsc.edu/~milanfar/software/sr-datasets.html.

[15] EPFL super resolution software, http://lcavwww.epfl.ch/software/superresolution/index.html.